Keystroke Biometrics for Student Authentication: A Case Study

Aythami Morales, Julian Fierrez
Universidad Autonoma de Madrid
Dept. de Tecnología Electronica y de las Comunicaciones, EPS, C\Francisco Tomas y Valiente, 28049 Madrid, Spain
{aythami.morales, julian.fierrez}@uam.es

ABSTRACT
This work presents a case study on the application of biometric systems for student authentication services. We analyze the accuracy of a keystroke dynamics algorithm used to authenticate students during a real online exam of an introductory computer science course.

Categories and Subject Descriptors
D.4.6 **[Security and Protection]:** *Access controls and Authentication*

Keywords
MOOCs, POOCs, authentication, keystroke dynamics, biometrics.

1. SUMMARY OF THE POSTER
The MOOCs have emerged in recent years as a new way to educate students in the basis of open, free, distributed and participatory courses. The online courses break with the barriers associated to traditional lessons and provide a new widely accessible education over the internet. However there is a discussion among all the academic sectors about the new educational scenario proposed by MOOCs. Among several topics, this controversial discussion includes the challenges related with the "certificates" or "statements of completion" of courses without classroom attendance. How can we certify that the student who is being certified is the one who perform the activities of the course? How can we avoid/detect anomalous users? The widely and accessible nature of the MOOCs increase the vulnerability of the systems and the authentication of the students emerges as a challenging and unsolved task. Researchers and MOOCs instructors are aware of the importance of reliable authentication for the future of online education and they have made efforts to propose and analyze convenient authentication approaches [1]. Among all the proposed approaches, biometric technologies seem to be the most attractive solutions. Biometric recognition technologies allow to authenticate users based on "something that users are" instead of traditional authentication based on "something that users know" such as PIN or passwords. Keystroke dynamic authentication is a well-known technology which has attracted much interest of industry and researchers during the last decade. Keystroke dynamics are interesting for MOOCs authentication because it is: i) transparent (it runs in the background without requiring explicit user interaction), and ii) continuous (authentication can be performed over all the user activity, not only based on an initial access to the platform). Transparent authentication is strongly related with the user experience and usability of the platform. The authentication service must be simple and not affect the normal activity of the MOOC. Continuous authentication is critical because an authentication system based on PIN codes assumes that the user will not provide his/her code to any other user. That assumption is not valid in a scenario in which the user could be interested in providing his/her PIN code to other user to pass an exam.

ITiCSE'15, July 6–8, 2015, Vilnius, Lithuania.
ACM 978-1-4503-3440-2/15/07.
http://dx.doi.org/10.1145/2729094.2754847

The experiments reported in this poster are performed over OhKBIC dataset (available at http://biometric-competitions.com/mod/competition/dataset.php?id=7). The database includes the responses of 64 students to five questions directly into the web-platform of the course which logged the keystroke pattern of each of the students (500 keystrokes per user). Therefore we consider here a text independent authentication scenario. The aim of the experiment is to analyze the performance of keystroke dynamics among the different responses. The experimental protocol is summarized as: i) for each user, the database is divided into enrollment (first 300 characters) and test (last 100 characters) not using the 100 characters in between in order to give time separation between enrollment and test data; ii) we search for common digraphs and trigraphs (sequences of two and three characters respectively) between the gallery and test set; iii) a cross-validation protocol is applied to measure the distance between gallery and test using a keystroke classifier based on Normalized Manhattan Distance [2]; iv) the final score is obtained as the mean of the best 40 distances (20 digraphs and 20 trigraphs); v) the performance of the overall experiment is provided according the average accuracy in authenticating each user (100-Equal Error Rate), see Table 1.

Table 1. Student authentication accuracy (100-EER)

Digraph	Trigraph	Combined
90.54%	91.97%	93.93%

The results show a promising performance with a correct student authentication over 90% using only 100 keystrokes. Although there is room for improvements, this work encourages to further explore in authentication services based on keystroke dynamics for online courses.

REFERENCES
[1] Miguel, J., Caballe, S., Prieto, J. 2013. Providing information security to MOOC: Towards effective student ´ authentication. In *Proc. of the Int. Conf. on Intelligent Networking and Collaborative Systems* (Xian, China), IEEE Press, 289–292.

[2] Morales, A., Fierrez, J., Ortega-Garcia, J. 2014. Towards predicting good users for biometric recognition based on keystroke dynamics. *In Proc. of Int. Workshop on Soft Biometrics*, 1-14, Zurich, Switzerland, LNCS 8926, 1-14.